图书情报工作
第 63 卷 第 4 期 2019 年 2 月

科学数据用户相关性标准研究*

张贵兰[1] 王健[1,2] 周国民[1,2] 刘建平[1] 韦草原[1]
[1]中国农业科学院农业信息研究所 北京 100081 [2]农业农村部农业大数据重点实验室 北京 100081

摘要：[目的/意义]以透镜理论为依据，从认知加工的角度出发，研究科学数据用户如何在数据共享平台中选取合适的数据。[方法/过程]研究分两个阶段进行，第一阶段选取 14 位被试通过半结构化访谈初步获取科学数据相关性标准集合及其使用情况；第二阶段通过发放 671 份问卷获取相关性标准的重要性，并对第一阶段获取的相关性标准内涵进行验证。[结果/结论]最终得到 9 个科学数据相关性标准，分别为主题性、可获得性、全面性、时效性、权威性、质量、便利性、规范性和可用性，并对其内涵进行了界定和验证。研究结果发现，全面性和规范性是科学数据的新增标准；可获得性、可用性和便利性存在很强的关联性；质量和规范性存在很强的关联性；质量与权威性虽然不相关，却保持一致的判断趋势。在未来的研究中为真正提升检索效率，改进检索系统，除考虑用户经常使用的标准之外，还要发掘那些使用频率不高但是很重要的标准。

关键词：科学数据 信息载体 相关性 相关性标准
分类号：G250
DOI：10.13266/j.issn.0252-3116.2019.04.014

1 引言

在生活、学习或者工作中，经常需要查询资料解决问题，因此就遇到一个问题：如何选取相关信息或资料。针对这个问题，已经有很多学者开展了研究。同样，在当今的数据时代，科学数据成为了日常工作中必不可少的分析材料，主要指通过科技活动或其他方式所获取到的反映客观世界的本质、特征、变化、规律等的原始基本数据以及根据不同科技需要进行系统加工整理的各类数据集。例如国家的关于各行各业的统计数据、卫星遥感数据以及基因数据等。这些数据中包含大量的信息，成为了主要信息载体之一。随着科学数据的发展，科学数据用户依然会遇到如何选取合适的数据的问题。

从用户的角度来说，科学数据用户在检索数据过程中，眼睛受到外界信息的刺激，将信息传递到大脑进行加工，最终做出相关性判断。在相关性判断过程中，信息加工就是在大脑中运用对应的相关性标准进行度量的过程。因此相关性标准是判断相关性等级的度量工具，其相关性的复杂性决定了相关性标准的复杂性。L. Schamber 等[1]研究总结了过去 30 年间发表的文献，得出的基本结论是相关性是多维的，在很大程度上依赖于搜索者的信息洞察力及其信息需求状况；相关性是动态的，因为对事物的评价会随着时间而改变，依赖于搜索者在某个时间点对信息和信息需求之间关系的判断；同时相关性也是可测量的，如果研究是从搜索者的角度进行概念上和操作上的处理的话。不同维度下的相关性要选择与其相适应的相关性标准进行度量。在信息载体维度上，从传统的纸质文字型信息发展到如今的网页型信息、图像型信息、音乐型信息甚至电子商务型信息，信息载体已经呈现出多元化，而相关性标准也发生了很大改变。如网页型信息中需要对多样性或安全性进行判断；图像型信息中需要对构图或吸引性进行判断；音乐型信息更多地偏向个人喜好。信息载体本身的物理特质和存在形式会对相关性标准产生相应的影响。因此当信息载体发展为科学数据时，其相关性标准必然发生变化。本文的研究重点主

* 本文系中国农业科学院科技创新工程(项目编号:CAAS-ASTIP-2016-AII)和国家社会科学基金项目“科学数据用户相关性标准与使用模式研究”(项目编号:14BTQ056)研究成果之一。
作者简介：张贵兰(ORCID:0000-0002-9153-3579)，博士研究生；王健(ORCID:0000-0003-4958-7669)，研究员，博士，通讯作者，E-mail:wangjian01@caas.cn；周国民，研究员，博士；刘建平(ORCID:0000-0002-1817-5373)，博士研究生；韦草原(ORCID:0000-0001-5259-4187)，数据分析师，硕士。
收稿日期:2018-07-03 修回日期:2018-09-17 本文起止页码:112-121 本文责任编辑:易飞

要有:①明确科学数据相关性标准;②科学数据相关性标准的内涵及其使用情况。

2 文献综述

一直以来,相关性的概念在情报学和信息检索领域都是一个难题。通过长时间的研究,学者们公认相关性是一个多维的、动态的,但同时又是一个系统化和可测量的概念。L. Schamber 等[1]强调了信息系统和检索中情境和情景的重要性。情景相关更注重查询对象的有效性。如在发表论文过程中,相关性至少是与论文写作不同阶段作者所处的情景相关,而且通常也与检索者在情景中的状态有关。因此学者 P. Ingwersen[2]提出相关性涉及的一个很重要的问题:是什么因素使一篇文献在第一次遇到时就呈现出相关,或是在更长的时间里才表现出相关。这是一个高度主观的相关性因素。本文将这类影响因素称为相关性标准,很多学者为回答这个问题耗费大量精力开展研究。

2.1 文献相关性方面

L. Schamber[3]从认知的角度在一个多媒体专业的工作环境中开展实验,研究了相关性标准。后续很多学者都沿用了这种实验方法开展研究。T. K. Park[4]、C. L. Barry[5]、W. Bruce[6]和 P. L. Wang[7]都试图通过实验的方法分解并定义相关性判断标准以及影响相关性判断过程的因素。L. Schamber 等[8]后续研究的重点是让标准和标准集更加规范和明确。她发表了一个包括 80 个相关性标准的汇编列表,这些标准是她总结前人相关文献后提出来的,并将其分为了 5 个大类。C. L. Barry 和 L. Schamber[9]发现不同情景下用户对相关性标准的选择有着很多相同之处,因此针对不同情境和用户下的两项研究进行对比,其结果发现不同维度的相关性标准研究中有 10 个相同相关性标准。剩余少数不同的标准主要是由于工作情境、搜索情境以及研究要求的差异。P. Borlund[10]在研究中也证实了这一发现:在不同工作环境的两组不同的用户共用了同一套相关性标准。这也初步证明了存在一套有限范围内普适性的标准集合,这类普适性的标准集合适用于各类用户、任务背景以及目标信息类型等。

2.2 网页相关性方面

进入 21 世纪以后互联网开始普及,人们开始在网页中寻找自己想要的信息。学者们开始研究如何设计系统或改进网页能让用户更快地从大量的网页信息中查找到自己所需要的信息。C. Silverstein 等[11]发现网络用户与传统信息检索系统中检索者的行为是不同的。S. Y. Rieh[12]通过实验和访谈相结合的方法得到了 3 个网页相关性标准:主题性、信息质量和认知上的权威性。A. Tombros 等[13]在 ACM 会议中探讨了网民在利用网页进行信息查询过程中进行相关性判断时用到的标准。A. Crystal 和 J. Greenberg[14]从不同于 A. Tombros 等[14]的角度研究了用户在进行网页搜索过程中对相关性标准的使用,研究选取了 12 个对健康信息感兴趣的用户开展研究,得到 8 个标准集,其中包括 22 个次级标准。R. Savolainen 等[15]在研究中提出了可靠性、安全性、多样性等新标准。Y. Kammerer 等[16]认为由于网页的特殊性,除了考虑内容信息,还需要考虑其质量和网页的类型。

2.3 图像相关性方面

信息化和数字化的发展带动了数字图像的发展,针对图像的相关性研究也越来越多。M. Markkula 和 E. Sormunen[17]首次对记者关于图像相关性标准的选取进行了研究,并得出 7 个相关性标准,其中技术性和视觉影响都是与图像本身特质相关的。Y. Choi 等[18]做了关于美国历史图像检索标准选取的研究,通过访谈分析共得到 9 个相关性标准,其中图像的吸引度和技术性都是图像所特有的标准。随着图像相关性标准研究的深入,T. Y. Hung 等[19]在研究中设定了不同的任务,最终得到 12 个图像相关性标准,其中美观感、构图、图片情节、外观等都与图像物理结构相关。S. Sedghi 等[20]在研究中采用扎根理论的方法,将图像与文档的相关性标准进行了对比分析,发现时效性、可用性、版权、颜色、目标观众、可靠性、信息大小、质量和主题性同样适用于文献检索;而方向、技术信息、放大率、原始性都是图像所特有的相关性标准。

通过上述回顾可以发现:①存在跨目标信息类型的共性标准,不同信息载体中的相关性标准存在着一定关联而非独立存在,如主题性、新颖性、时效性、质量等;②文本相关性标准研究是基础性研究,其他研究都是在其基础上开展的,因为文本是最主要也是使用最多的信息载体;③任何信息载体都会存在与其本身特定物理性质相关的特异性标准,如网页信息载体类型下会关注超链接,还会有学者提及网页的安全性问题,图像信息载体类型下会出现技术性和吸引度等相关性标准;④学者们针对各类信息载体都开展了研究,但是科学数据作为新型信息载体,还无人研究。

因此本文开展以科学数据为信息载体的相关性标准研究。在研究过程中,既要考虑其文本信息固有的

相关性标准,又要考虑科学数据本身特有性质产生的新型相关性标准。

3　理论框架及实验设计

3.1　理论框架

美国心理学家希伦斯维克提出我们的感官并不能与外部世界的物体和事件发生直接的联系,而只能通过介于外部事物和内部直觉之间的“透镜”来获取信息[21]。如图 1 所示,模型左侧是人所处的真实世界的事件和事物,中间是人做出判断时头脑中的心理过程,右侧是人根据判断做出的相应的行为。依据透镜模型,本文探索如何将感官刺激和用户做出判断之间的内在联系起来,并通过线性组合模拟用户做出判断。

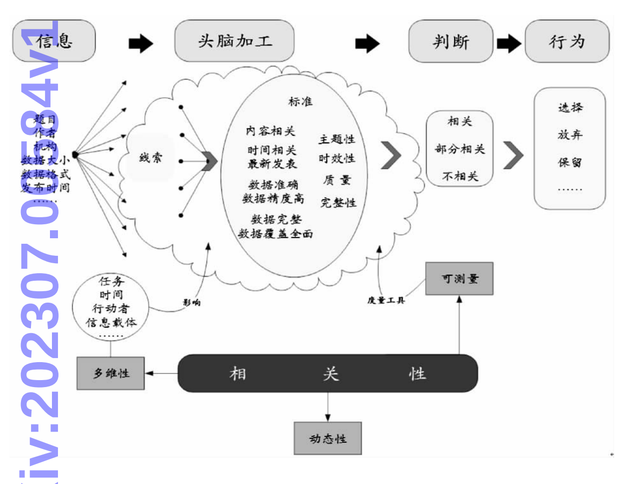


图 1　理论模型

从传统意义上讲,对信息的判断和选择基于相关性判断的模型,即用户接受或放弃某条信息的主要依据在于这条信息是否与用户的需求相关。相关性判断隐含在用户的决策之中[22]。根据决策理论,从一系列检索结果中进行选择并做出最后的决策判断都要依靠决策者所使用的相关性标准[23]。学者们研究发现用户在判断相关性的过程中除了主题相关标准外,还会使用其他多样化的标准来进行判断。

受信息论的启发,认知心理学家将人视为能够接受、处理和加工信息的传输装置[24]。人具有对信息进行串行和并行加工的能力。人可以被视为一个通信通道,人能同时传递的信息量是有限的,但能通过信息编码克服通道容量的局限。因此相关性判断在认知中是一个系列加工的过程,进而相关性标准的使用也是一个系列加工的过程。

3.2　实验设计

本研究使用的研究方法有访谈、情景再现刺激用户认知、调查问卷。数据分析方法主要是 Nvivo 内容分析和利用 SPSS 软件进行定量分析。整个实验分为两阶段,第一阶段为访谈,研究选取了被试 14 人,来自全国各地的大学以及研究机构,如中国科学院地理研究所、北京航空航天大学、中国地质大学等,其中博士 1 人,硕士 11 人,本科 2 人。用户再现完成“共享杯”竞赛项目 时的检索过程,笔者针对检索过程对其进行了访谈,初步获取科学数据相关性标准集合以及用户对其内涵的界定。第二阶段的实验针对“共享杯”参赛学生,发布大量实验问卷,验证第一阶段的访谈结果。

3.2.1　访谈数据收集和处理　为了获取用户完成任务时头脑中所使用的相关性标准,实验采用半结构化访谈。对访谈内容进行编码,用 Nvivo11.0 进行内容分析,获取用户在检索科学数据过程中所使用的标准以及对标准内涵的界定。

相关性标准是用户头脑中存储的知识,在访谈中体现为工具性的概念或逻辑关系[25],如“与我的研究相关”“看看能不能用”“信息非常全面”等,将这些概念性的词语编码为标准。结合前人的研究和定义,根据其意义命名,分别为“主题性”“可用性”“全面性”(见表 1)。

表 1　编码实例

访谈	标准
IP:最主要的首先看一下这个质量怎么样,其次能不能用,能不能为我所用。从这三个方面来把握。 I:那你这个为你所用是指什么? IP:跟论文相关。	质量 可用性 主题性
IP:就是数据生产的那个单位,中科院可信度比较高,所以我就比较放心,其他学校获取的数据就不太敢用	权威性
IP: Modies 下载的话,比较好下,国内国外都有	便利性

3.2.2　问卷数据收集和处理　实验共计收到 671 份问卷,剩余有效问卷 544 份,有效率为 81.07%。其中本科生 22 人,硕士研究生 421 人,博士研究生 99 人,博士后 2 人。调研对象在平时的工作和学习中经常(50% 的工作)使用科学数据的人群所占比重为 68%,有时(20% –50% 的工作)会使用科学数据的人群占 23.9%,所以被试代表性强,具有一定的参考价值。

4　结果

4.1　科学数据相关性标准

通过编码,共得到 9 个科学数据相关性标准(见表 2)。分别是主题性、可获得性、质量、规范性、权威性、全面性、便利性、可用性和时效性。

表2 标准定义

标准	定义
主题性	数据与用户的研究相符合，如数据内容、数据涉及时间、数据涉及区域
可获得性	不受外界因素影响，用户可以顺利获取数据。如：没有获取权限、无下载链接、价格高昂等
质量	数据的优劣程度，如数据是否准确、正确或有效
规范性	数据的分类体系，采集过程等是否符合国家要求，或是否一致
权威性	用户可以信任这条数据，主要指发布数据的某人或某个机构是极具影响力的
全面性	数据覆盖全面，或是完整没有缺失
便利性	数据的检索、获取、使用过程方便
可用性	数据可以使用，不会受认知局限或格式等问题而导致不能或不会使用
时效性	数据仅在一定时间段内对研究具有价值，如：数据的发表时间是最新发表的，或没有过时的

4.1.1 主题性 主题性被定义为数据与用户的研究相符合，主要包含3个方面：内容相关、时间相关、区域相关。科学数据本身包含一定的内容信息，对于农业数据来说，可能包含产量、耕地面积等；除此之外还会包含相应的时间信息和空间信息，如某年某地的产量、某天某地的降雨量等。因此用户在选择科学数据的过程中，不仅要选择研究内容相关的，还有选择符合自己的研究时间段和研究区域的数据。编码为主题性的例子有："因为数据很多，根据时间和区域，地震发生时间是2014年6月份，我们找的是2014年5月、6月、7月、8月没有云的影像数据""遥感数据是非常大的，要做一个小区域就会找对应区域经纬度，要做全国的就会找全国的数据""他给的数据就是很多年土地数据和经济数据，跟题目还挺相关的。"

4.1.2 可获得性 可获得性是指不受外界因素影响，用户可以顺利获取数据。得到相关数据是用户检索的最终目的，如果不能获取则数据就是无用的。正如R. Savolainen等[15]在研究中提到不可得到会让用户直接否决其信息。通过访谈可以发现，影响用户、使其无法获取数据的原因有没有获取权限、没有下载链接、价格昂贵等。如用户提到："还有就是资金问题，我们尽可能地会选择一些免费的数据，不会买质量特别好、特别贵的数据""唯一的遗憾是下载不下来，无法全文阅读。"

4.1.3 质量 质量是指数据的优劣程度。S. Y. Rieh[12]认为"有用性"和"优秀的"是信息质量中两个最基本的因素。P. L. Wang[7]在研究文献相关性标准时也对质量下了定义：文献是优秀的。优秀是一个宽泛的概念，无法准确说明如何评价质量的优劣。广义的质量定义可能要包含数据的各个方面，也许还会有一个综合的评价指标；而在本文中数据质量就选取其最主要的内涵，即数据是准确的、正确的、有效的。编码为质量的例子有："目前在网上降水量没有一个比较准确的数字""从很多细致的统计途径得到的数据""一个是卫星获取数据，另一个是人工到野外获取的数据，人工获取的精度肯定是比卫星拍摄的好。"

4.1.4 规范性 规范性是指数据的分类体系、采集过程等是否符合国家要求，或是否一致。科学数据作为科研工作的基础数据，是为科研工作服务的。采集过程、分类标准的规范性直接影响着工作的进度。例如有用户提到"发现数据没办法整理，十年的数据，每年的处理都不一样，不太好归纳整理。""他们统计方法是采用国家标准或者国际标准，这些标准对于长时间序列数据来说是很规范的。"

4.1.5 权威性 权威性主要指发布数据的某人或某个机构是极具影响力的。发布者或发布机构的权威性使得用户信任某条数据。C. Watson[23]认为认知权威性影响一个人的思想，使其有意识地辨认一篇文章。也就是说权威性具有很强的个人主观色彩，对权威性评价的高低也会因人而易。正如用户提及的"最后会看数据来源，看到是中科院，我会比较放心""国家花那么多人力物力统计经济数据然后权威发布，数据的质量和可信度会更好。"

4.1.6 全面性 全面性是指数据覆盖全面，或完整没有缺失。在科研工作中需要的科研数据可能是连续很多年的，或者综合各项指标，中间缺失某一年或某项指标都不能完成任务。编码为全面性的有："我们现在用的数据集比较多，持续时间长，所以看一下时间序列，是什么时候到什么时候，中间有没有缺，是否完整""看指标的完整程度，指标完整就采纳，指标不完整就删去。"

4.1.7 便利性 便利性是指数据的检索、获取、使用过程方便。用户在获取数据的过程中，某条数据的付出成本很小，会让用户更加青睐——即最小努力原则。面对信息，人总是有意识地回避复杂、费事和不熟悉的信息源，而去寻找快捷方便的信息途径。语言学家G. K. Zipf[26]1949年的研究中就说明人们在进行信息搜索时会采用阻力最小的途径。编码为便利性的有："Modies下载的话，比较好下，国内国外都有""因为我们在地理所的申请是比较快的，不需要离线申请，只需要提交一个表格就够了。"

4.1.8 可用性 可用性是指在获取数据之后，数据可

以使用，不会受认知局限或格式等问题而导致不能或不会使用。科学数据本身格式类型多样化，专业要求高，且受领域认知局限较大。所以数据的可用性对于科研工作也至关重要。编码为可用性的有：“这条 landset 数据，虽然时间地点吻合，但是云量特别大，不能用”“我是学地理的，要看格式是否有问题。”

4.1.9　时效性　时效性是指数据仅在一定时间段内对研究具有价值，过了某个时间段，数据就没有了价值。同一条数据在不同的时间具有很大的性质上的差异，这个差异性叫时效性，时效性影响数据的生效时间，决定了数据在哪些时间内有效。一旦过了这段时间，数据就失去了存在的价值，或者说价值大打折扣。编码为时效性的有：“因为土地利用也会有变化呀，最新的才会接近现在的情况”“如果还能从一些部门获得更实时的一些数据就好啦。”

通过对 14 个用户的访谈，共计得到 518 条回应，用户总共检索数据 345 条。通过 518 回应的编码，各项标准共计被提及 287 次。表 3 展示了各项标准被提及的频次，所占百分比以及提及被试的个数。通过表 3 可以看出，使用最为频繁的标准为主题性，其次为质量。主题性、质量和可获得性的使用频率之和在 70% 左右，少数相关性标准被频繁地使用，接下来会有大部分偶尔使用的标准。

表 3　标准频次

标准	频次	百分比	被试个数
主题性	106	36.9%	14
质量	63	21.9%	11
可获得性	33	11.5%	11
时效性	22	7.7%	8
全面性	20	7.0%	10
便利性	19	6.6%	9
可用性	11	3.8%	8
权威性	7	2.5%	4
规范性	6	2.1%	2
总计	287	100%	–

4.2　科学数据相关性标准使用及内涵

通过问卷调查，我们获取了 9 个科学数据相关性标准的重要程度分布图（见图 2），从图 1 中可以看出，最重要的标准为质量，其次为主题性，最不重要的标准为便利性。质量（4.18）与便利性（3.48）之间相差仅 0.7 分，由此可见在被试者心中，9 个相关性标准重要程度差距不大，他们认为每个独立的标准对相关性判断都比较重要；但是在研究相关性标准使用过程中，我们会发现一些标准被频繁地使用，如主题性、质量；还有一些标准几乎不被使用，如权威性。由此看来有些标准不经常使用，但是不代表不重要。使用频率和重要程度并不呈正相关。

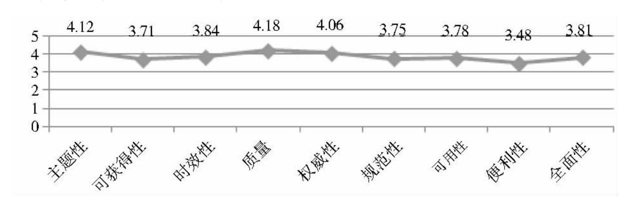


图 2　相关性标准重要程度分布

因此，本研究将相关性标准的使用频率和重要性依据排位进行了对比（见表 4），排名差距最大的是可获得性和权威性，其次是便利性。被试者在真实检索过程中，很大程度会考虑数据是否可以获取以及获取过程中付出的成本，而在进行重要性排序时，却认为两者不重要；同样被试者认为权威性是重要的，而在检索过程中却很少用到。由此可见，被试者在认知层面上，清楚地知道哪些标准可以帮助获得“最优”的数据；但是在真正的检索过程中，由于人类的惰性或外界环境的限制，被试选择数据的依据是“速度”而非“最优”，即希望通过最小努力获取最大收益。

表 4　相关性标准使用频率与重要性对比

排名	第一位	第二位	第三位	第四位	第五位	第六位	第七位	第八位	第九位
使用频率	主题性	质量	可获得性	时效性	全面性	便利性	可用性	权威性	规范性
重要性	质量	主题性	权威性	时效性	全面性	可用性	规范性	可获得性	便利性

为了验证第一阶段实验相关性标准内涵的准确性，将访谈过程中出现的关于各项标准的描述性语句总结为短语，让 544 位被试进行选择归类。每一个描述性的短语都可以自由归类，也可以同时归类到不同的标准下。统计将短语归为同一标准下的人数，可以得出一个雷达图（见图 3）。通过总结，得到了被试一致性较高的 9 个相关性标准的描述性短语（见表 5）。

通过表 5 可以看出，主题性、全面性和时效性的描述性短语没有交叉混淆。但是被试者只认为内容相关为主题性，而时间相关为时效性，且空间相关不显著。由此说明两点：一是被试选择主要依据字面意思，没有深入思考短语含义；二是调查问卷的表述还不够简单直观明了。

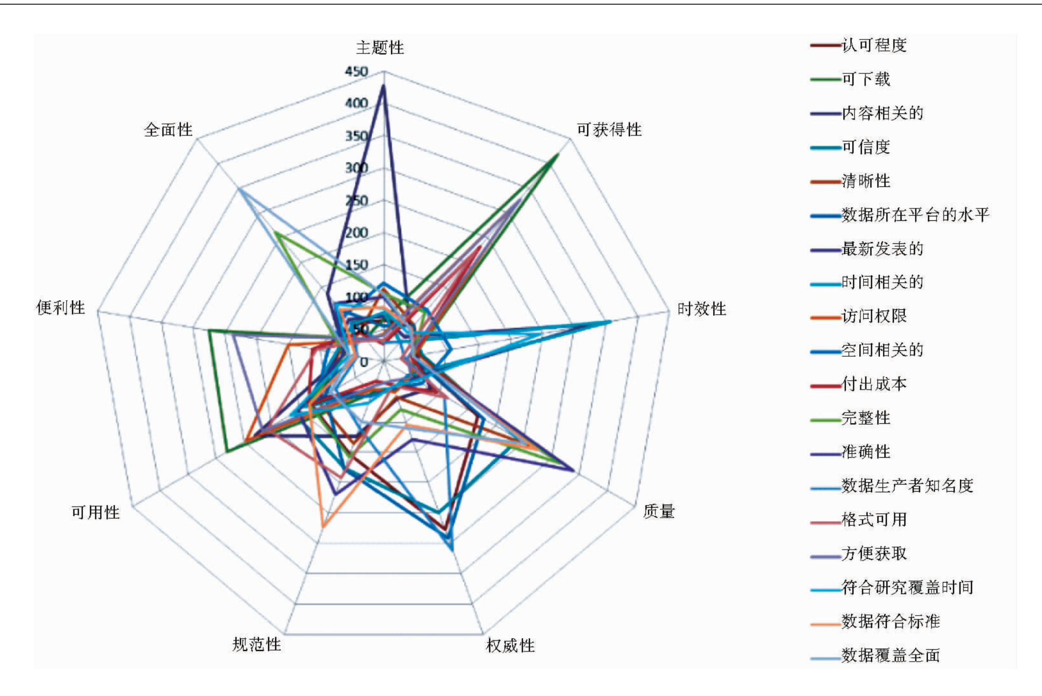


图3　相关性标准分类雷达图

表5　描述性短语归类

标准	描述性短语
主题性	内容相关的
全面性	完整性、数据覆盖全面的
时效性	符合研究覆盖时间的、最新发表的、时间相关的
质量	清晰性、可信度、准确性、数据符合标准
权威性	数据所在平台、认可程度、可信度、数据生产者知名度
规范性	数据符合标准
可获得性	可下载、访问权限、付出成本、方便获取、格式可用
可用性	格式可用、方便获取
便利性	方便获取

针对质量、权威性和规范性，我们发现可信度既属于质量，又属于权威性；数据符合标准既属于质量又属于规范性。由此，本研究团队需要思考两点：一是规范性是否属于评价质量的一个分支，就像准确性和清晰性一样，且对质量和规范性进行皮尔逊相关性检验，得出显著性 P＝0.000，在 α＝0.01 水平下显著相关；二是质量和权威性是否存在交叉（见图4），且对质量和权威性进行皮尔逊相关性检验，得出显著性 P＝0.162，由此可见两者并不相关，但是从图4可以清晰地看出两者在某些描述性短语选择的趋势上保持一致，说明两者有着本质区别，但却又有着一定的关联性，如是否数据越权威质量越好。

针对可获得性、可用性和便利性，可以看出可用性包含便利性的短语（方便获取），可获得性又全部包含了可用性的短语（格式可用、方便获取）。从三者的对比分析图（见图5）中可以看出三者的趋势基本保持一致。通过对三者分别进行皮尔逊相关性检验，显著性 P 均小于0.01，可见三者均显著相关，且便利性和可获得性的皮尔逊指数高达0.932。可见三者息息相关，只有获得数据才能进一步判断是否可用，而在此过程中便利性则决定了获取数据过程中的成本问题。

5　讨论

本项研究主要针对的信息载体为科学数据，并得出了9个科学数据相关性标准，分别为主题性、可获得性、全面性、时效性、权威性、质量、规范性、便利性和可用性。其中主题性、可获得性、时效性、权威性、质量、便利性、可用性是与之前很多学者（C. L. Barry[5]，L. Schamber[8]，P. Wang[27]等）在不同信息载体下的研究相一致的标准。除此之外，对于主题性，后续很多学者（A. Crystal 等[14]，R. A. Hamid[28]等）的研究中也有“主题性”，且使用频率也是最高的，S. G. Hirsh[29]在针对学生选取“electronic resources”的研究中发现，主题性和个人兴趣是最重要的相关性标准。对于可获得性，M. Markkula 等[17]在研究中表明当检索结果可以获得时，记者们得到了令人满意的结果，S. Sedghi 等[30]也认为可获得性对于用户来说是很重要的。对于时效性，C. Papaeconomou 等[31]和 P. Balatsoukas 等[32]在研究网页相关性标准时也提出了时效性，Y. Choi 等[33]认为时间范围对图像的选择也有一定的影响。对于权威性和质量，这是判断信息真实性和可靠性的重要标准，也是得到学者们一致认可的相关性标

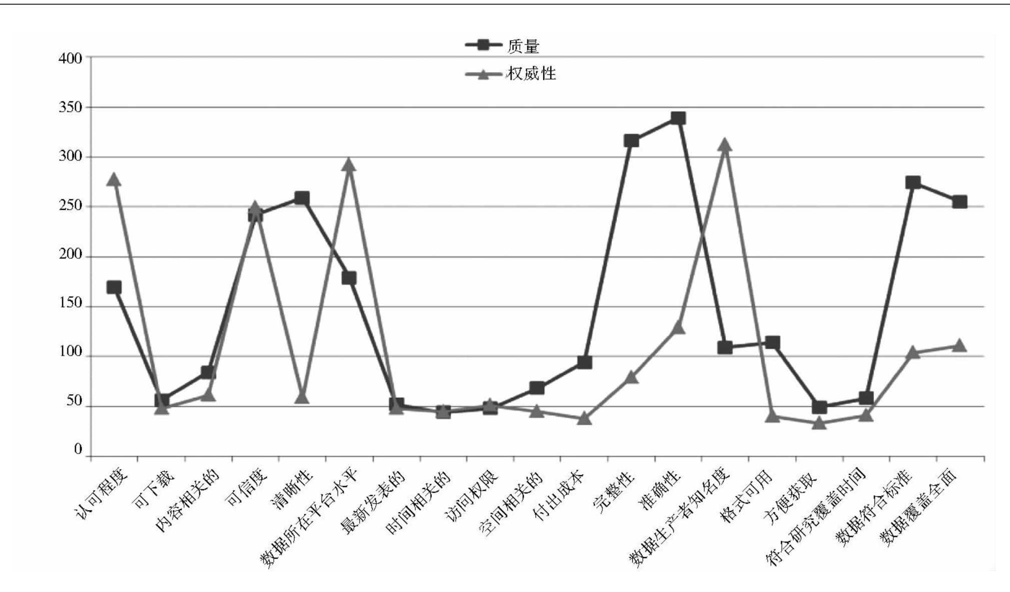


图 4　质量和权威性的对比分析

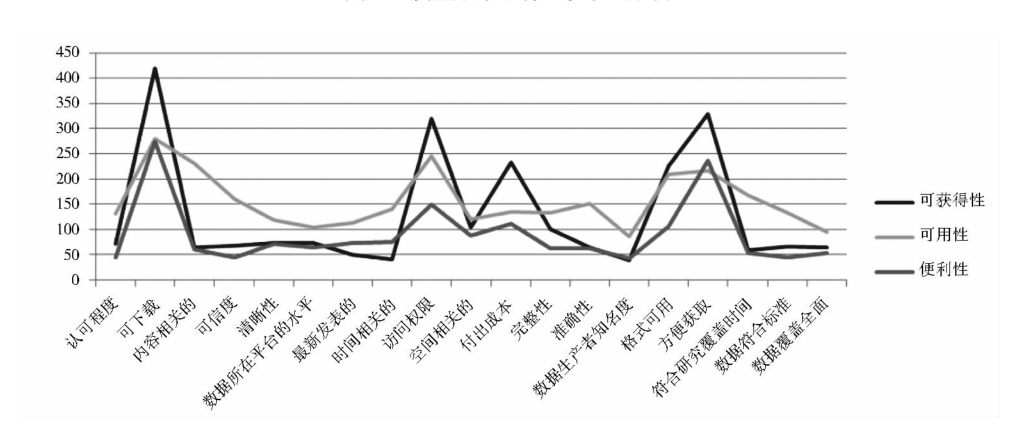


图 5　可获得性、可用性和便利性的对分分析

准，而且 S. Y. Rieh[12] 专门针对两者对信息判断的影响开展研究。对于便利性，主要指付出成本的大小，面对信息，人总是有意识地回避复杂、费事和不熟悉的信息源，而去寻找快捷方便的信息途径，M. D. Barnes 等[34] 在针对用户在网络中搜索健康信息时所使用的相关性标准的研究中就提到了使用的便利性，且早年间诸多研究都表明信息获取付出成本越小，越容易被使用[29-31]。对于可用性，A. Crystal 等[14] 则提出格式是否可用，R. Savolainen 等[15] 则提出语言是否可以理解可用。由此可见以上 7 个科学数据相关性标准与前人的研究是一致的，即是跨越了不同的信息载体类型而存在的。这 7 个相关性标准适用于所有的信息类型，如文档、图像、音乐、视频、数据等。

与此同时，本研究还发现了与科学数据本身特质息息相关的独特相关性标准，即全面性和规范性。首先，在进行前人的研究综述中，没有学者在文档、图像、网页、音乐等信息载体的研究中提出全面性和规范性两个标准；其次，本文通过对访谈的提炼总结可以发现，全面性主要侧重在数据在时间和区域序列上的连续性和完整性，这是基于数据的定时性和定位性的；最后，规范性侧重数据的分类体系和统计方法方面符合标准，这是因为科学数据的检索是基于数据的共享，而若国内外缺少统一的科学数据标准规范体系，则妨碍了不同领域科学数据的交换与共享[38]，从而影响科学数据的检索，科学数据具有极强的专业性和领域性，只有统一规范，才能更好地实现数据的共享。A. Laplante[39] 在关于音乐相关性标准研究的论文中曾提出，虽然研究发现一些在文档网页中出现的相关性标准依然适用音乐环境（如质量和权威性等），但是还是会出现一些独特的相关性标准。正如 T. Saracevic 等指出：相关性的研究不能脱离情景，而且还要考虑到情景内在因素和外在因素动态交互的影响[40]。

通过文献总结可以发现，很多学者在研究中都提及了新颖性。C. L. Barry[5] 曾提出了内容新颖性和信

息来源新颖性标准;P. L. Wang[7]在研究中也提出在不考虑出版时间的前提下,信息内容对用户来说是否是新的影响用户判断的因素;A. Tombros等[13]提到在网页检索中网页是否已经看过也会影响用户的判断;R. Savolainen等[15]指出了好奇心和熟悉度对用户的影响;M. Twait[41]也在研究中提到了熟悉度对用户的影响;C. Papaeconomou等[29]、A. R. Taylor等[42]、R. A. Hamid等[28]在研究相关性标准的使用过程中也都选择了新颖性。然而本文中并没有出现新颖性,推测原因主要是:科学数据实用性较强,用户在检索之前就已经明确了所需要的数据,以用为主,并不需要获取灵感或是启发。

在相关性标准使用方面,使用最为频繁的标准为主题性,其次为质量,这与之前很多学者的研究都一致(P. Balatsoukas等[32]、S. Sedghi等[30]、R. A. Hamid等[28])。主题性、质量和可获得性的使用频率之和在70%左右,少数相关性标准被频繁地使用,接下来会有大部分偶尔使用的标准,这一发现在A. Crystal等[14]的研究中也被证实过,该发现符合著名的长尾定律。虽然权威性对于用户来说很重要,但几乎不被提及,A. Crystal等[14]在研究中也得到了同样的结论。其原因有两点:一是有些用户不能很好地区分质量和权威性,或者将权威性归为质量,导致权威性使用频率较低;二是在初步满足内容相关和质量要求以后,用户会以成本最小为原则选择。付出成本相同的情况下,才会考虑权威性的等级。由此可见使用频率与重要程度不成正比,A. Tombros[43]和C. L. Barry等[9]在早年的研究中也提及此点。

对于相关性标准内涵的界定,早在1996年,L. Schamber等[8]指出受到情景和被试人群的影响,学者们提出了不同的层级和内涵的相关性标准。经过多年的研究,有些标准的内涵已经十分明了,但是有些标准内涵依然模糊,如质量。本文通过对前人研究的总结和对访谈内容的提炼,将质量定义为数据是准确的、正确的、有效的。而A. R. Taylor等[42]认为质量包含以下5个“价值”:准确性、全面性、时效性、可靠性和有效性。J. E. Klobas[44]提出质量的4个成分为:准确性、权威性、时效性和新颖性。S. L. Caudle等[45]认为质量的概念应该包括实际价值、感知价值、审美、自身特征以及随着时间变化后的自身意义。S. Y. Rieh[12]则认为有用的和好的是信息质量中两个最基本的因素。由此可见就广义来说,学者们一致认为质量主要用来评价信息的优劣程度,但是具体的内涵却是仁者见仁,智者见智。

6 结论

通过本项研究,初步得到了以科学数据为信息载体的相关性标准集合,共计9个相关性标准。分别为主题性、可获得性、全面性、时效性、权威性、质量、便利性、规范性和可用性。其中主题性、可获得性、时效性、权威性、质量、便利性和可用性是跨越不同信息载体而存在的共性标准;全面性和规范性是与科学数据自身特质相关的独特标准。

对访谈过程中相关性标准出现频率的统计发现,主题性、质量和可获得性使用最为频繁;从用户对相关性标准重要程度的打分来看,质量、主题性和权威性最为重要。由此可见使用频率和重要程度并不成正相关。为了真正提升检索效率,改进检索系统,在考虑用户经常使用的标准之外,还要发掘那些使用频率不高但是很重要的标准,如权威性。

通过大量用户对相关性内涵的归类,已经初步得到了主题性、全面性、时效性的内涵;质量和规范性的内涵,便利性、可获得性和可用性的内涵高度相关,存在包含关系;质量和权威性的内涵并不相关,但是在某些描述性短语选择的趋势上保持一致,说明二者本质存在区别,但是依然存在某些特定的关联。

参考文献:

[1] SCHAMBER L, OTHERS A. A re-examination of relevance: toward a dynamic, situationaldefinition[J]. Information processing & management, 1990, 26(6):755－776.
[2] INGWERSEN P. Cognitive perspectives of information retrieval interaction: elements of a cognitive IR theory[J]. Journal of documentation, 1996, 51(1):3－50.
[3] SCHAMBER L. Users' criteria for evaluation in a multimedia environment[C]// Proceedings of the Association for Information Science (ASIS) 54th annual meeting. Washington:ASIS,1991:126－133.
[4] PARK T K. The nature of relevance in information retrieval: an empirical study[J]. Library quarterly, 1993, 63(3):318－351.
[5] BARRY C L. User-defined relevance criteria: an exploratory study[J]. Journal of the Association for Information Science & Technology, 2010, 45(3):149－159.
[6] BRUCE H W. A cognitive view of the situational dynamism of user-centered relevance estimation[J]. Journal of the Association for Information Science & Technology, 2010, 45(3):142－148.
[7] WANG P L. The design of document retrieval systems for academic users: implications of studies on users' relevance criteria[J]. Proceedings of the American Society for Information Science & Tech-

nology, 1997, 34:162 – 173.
[8] SCHAMBER L, BATEMAN J. User criteria in relevance evaluation: toward development of a measurement scale[C]// Proceedings of the ASIS annual meeting. Baltimore: ASIS, 1996:218 – 225.
[9] BARRY C L, SCHAMBER L. Users' criteria for relevance evaluation: a cross-situational comparison[J]. Information processing & management, 1998, 34 (2/3):219 – 236.
[10] BORLUND P. The concept of relevance in IR[J]. Journal of the American Society for Information Science and Technology, 2003, 54(10):913 – 925.
[11] SILVERSTEIN C, MARAIS H, HENZINGER M, et al. Analysis of a very large Web search engine query log[J]. ACM sigir forum, 1999, 33(1):6 – 12.
[12] RIEH S Y. Judgment of information quality and cognitive authority in the Web[J]. Journal of the Association for Information Science & Technology, 2010, 53(2):145 – 161.
[13] TOMBROS A, RUTHVEN I, JOSE J M. Searchers' criteria for assessing Web pages[C]// International ACM SIGIR conference on research and development in informaion retrieval. Toronto: ACM, 2003:385 – 386.
[14] CRYSTAL A, GREENBERG J. Relevance criteria identified by health information users during Web searches[J]. Journal of the American Society for Information Science & Technology, 2014, 57(10):1368 – 1382.
[15] SAVOLAINEN R, KARI J. User-defined relevance criteria in Web searching[J]. Journal of documentation, 2006, 62(6):685 – 707.
[16] KAMMERER Y, AMANN D G, GERJETS P. When adults without university education search the internet for health information: the roles of Internet-specific epistemic beliefs and a source evaluation intervention[J]. Computers in human behavior, 2015, 48(C): 297 – 309.
[17] MARKKULA M, SORMUNEN E. End-user searching challenges indexing practices inthe digital newspaper photo archive[J]. Information retrieval, 2000, 1(4):259 – 285.
[18] CHOI Y, RASMUSSEN E M. Users' relevance criteria in image retrieval in American history[J]. Information processing & management, 2002, 38 (5): 695 – 726.
[19] HUNG T Y, ZOELLER C, LYON S. Relevance judgments for image retrieval in the field of journalism: a pilot study[C]//International conference on Asian digital libraries. Berlin: Springer, 2005: 72 – 80.
[20] SEDGHI S, SANDERSON M, CLOUGH P. How do health care professionals select medical images they need? [J]. Aslib proceedings new information perspectives, 2013, 65(1):54 – 72.
[21] 海斯蒂,道斯. 不确定世界的理性选择[M]. 北京:人民邮电出版社, 2013.
[22] SARACEVIC T. Relevance: a review of the literature and a framework for thinking on the notion in information science. part III: behavior and effects of relevance[J]. Journal of the Association for Information Science & Technology, 2014, 58(13):2126 – 2144.
[23] WATSON C. An exploratory study of secondary students' judgments of the relevance and reliability of information[J]. Journal of the Association for Information Science and Technology, 2014, 65(7):1385 – 1408.
[24] 武秀波. 认知科学概论[M]. 北京:科学出版社, 2007.
[25] 高飞, 石蕾, 王健,等. 农业科学数据用户相关性线索与标准之间关系探索[J]. 图书情报工作, 2017,61(15):72 – 80.
[26] ZIPF G K. Review of human behavior and the principle of least effort: an introduction to human ecology[M]. Cambridge: Addison-Wesley Press, 1949.
[27] WANG P, SOERGEL D. A cognitive model of document use during a research project. Study I. document selection[J]. Journal of the Association for Information Science & Technology, 1998, 49(2):115 – 133.
[28] HAMID R A, THOM J A, ISKANDAR D A. Effects of relevance criteria and subjective factors on Web image searching behaviour[J]. Journal of information science, 2017, 43(6):786 – 800.
[29] HIRSH S G. Children's relevance criteria and information seeking on electronic resources[J]. Journal of the American Society for information Science, 1999, 50(14): 1265 – 1283.
[30] SEDGHI S, SANDERSON M, Clough P. A study on the relevance criteria for medical images[J]. Pattern recognition letters, 2008, 29(15): 2046 – 2057.
[31] PAPAECONOMOU C, ZIJLEMA A F, INGWERSEN P. Searchers' relevance judgments and criteria in evaluating Web pages in a learning style perspective[C]// International symposium on information interaction in context. London: ACM, 2008:123 – 132.
[32] BALATSOUKAS P, RUTHVEN I. An eye - tracking approach to the analysis of relevance judgments on the Web: the case of Google search engine[J]. Journal of the American Society for Information Science and Technology, 2012, 63(9): 1728 – 1746.
[33] CHOI Y, RASMUSSEN E M. Users' relevance criteria in image retrieval in American history[J]. Information processing & management, 2002, 38(5): 695 – 726.
[34] BARNES M D, Penrod C., NEIGER B L, et al. Measuring the relevance of evaluation criteria among health information seekers on the Internet[J]. Journal health psychol, 2003, 8(1):71 – 82.
[35] CURLEY S P, CONNELLY D P, RICH E C. Physicians' use of medical knowledge resources: preliminary theoretical framework and findings[J]. Medical decision making, 1990, 10 (4):231 – 241.
[36] GERSTBERGER P G, ALLEN T J. Criteria used by research and development engineers in the selection of an information source [J]. Journal of applied psychology, 1968, 52 (4):272
[37] HARDY A P. The selection of channels when seeking information: cost/benefit vs least-effort[J]. Information processing & manage-

ment,1982, 18 (6):289－293.

[38] 伯格曼. 大数据、小数据、无数据:网络世界的数据学术[M]. 孟小峰等,译. 北京:机械工业出版社,2017

[39] LAPLANTE A. Users' relevance criteria in music retrieval in everyday life: an exploratory study[C]// International Society for Music Information Retrieval conference. Utrecht, Netherlands:DBLP, 2012:601－606.

[40] SARACEVIC T. Relevance: a review of the literature and a framework for thinking on the notion in information sicence. Part II: nature and manifestations of relevance[J]. Journal of the American Society for Information Science and Technology 2007, 58(13): 1915－1933

[41] TWAIT M. Undergraduate students' source selection criteria: a qualitative study[J]. Journal of academic librarianship, 2005, 31(6):567－573.

[42] TAYLOR A R, COOL C,BELKIN N J. Relationships between categories of relevance criteria and stage in task completion[J]. Information processing & management,2007, 43(4):1071－1084.

[43] TOMBROS A, RUTHVEN I, JOSE J M. How users assess Web pages for information seeking[J]. Journal of the Association for Information Science & Technology, 2014, 56(4):327－344.

[44] KLOBAS J E. Beyond information quality: fitness for purpose and electronic information resource use[J]. Journal of information science, 1995, 21(2):95－114.

[45] CAUDLE S L, MARCHAND D. A. Managing information resources: new directions in state government[J]. Information management review, 1990, 166(5):9－30.

作者贡献说明:

张贵兰:负责实验设计,分析数据,起草、撰写并修改论文;

王健:负责指导设计实验,提出研究框架,指导论文修改;

周国民:提出论文修改意见;

刘建平:协助完成实验和后期数据分析处理;

韦草原:负责数据的采集与整理。

User-defined Relevance Criteria in Scientific Data

Zhang Guilan[1] Wang Jian[1,2] Zhou Guomin[1,2] Liu Jianping[1] Wei Caoyuan[1]

[1] Agricultural Information Institute of Chinese Academy of Agricultual Sciences, Beijing 100081

[2] Key Laboratory of Agricultural Big Data, Ministry of Agriculture, Beijing 100081

Abstract: [**Purpose/significance**] Based on lens theory and from the perspective of cognitive processing, the paper studied how the users of scientific data select relevant data from the data sharing platform. [**Method/process**] The study was conducted in two stages. In the first stage, a total of 14 subjects were selected to obtain their relevance criteria and usage of scientific data through semi-structured interviews. The level of importance of the relevance criteria was determined, and in the second stage, the concepts of the relevance criteria obtained in the first stage were further verified through 671 questionnaires. [**Result/conclusion**] Finally, 9 relevance criteria were determined for scientific data: topicality, availability, comprehensiveness, currency, authority, quality, convenience, standardization, and usability, and the defined these concepts. The results showed that comprehensiveness and standardization are unique criteria related to the nature of scientific data. The concepts of convenience, availability, and usability are highly associated. The concepts of quality and authority are irrelevant, but they are consistent in some descriptive phrases. Thus, the concepts that define them must be further clarified. In order to truly make a better data search engine and improve its search efficiency, moving beyond the criteria often used by users, it is necessary to determine those criteria that are not often used, but still very important.

Keywords: scientific data information carrier relevance relevance criteria